

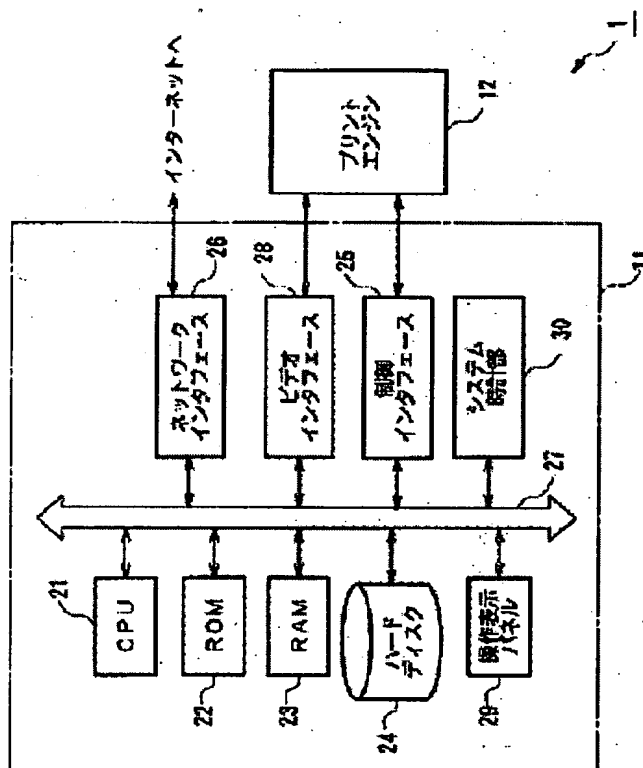
**DOCUMENT PROCESSOR**

**Patent number:** JP2001350745  
**Publication date:** 2001-12-21  
**Inventor:** UCHIDA MASAKI  
**Applicant:** FUJI XEROX CO LTD  
**Classification:**  
- **International:** G06F17/21  
- **European:**  
**Application number:** JP20000172635 20000608  
**Priority number(s):**

**Abstract of JP2001350745**

**PROBLEM TO BE SOLVED:** To provide a document processor, which considers pull-print, capable of improving convenience by easily setting the conditions of recursive acquisition.

**SOLUTION:** The device has a means for holding a managing table relatively describing reference information, the conditions of recursive acquisition and the processing conditions of an acquired document, means for receiving the designation of the reference information, a means for recursively acquiring document data according to the conditions of recursive acquisition related to the relevant designated reference information, which becomes a starting point, on the basis of reference information set in document data to refer to inside the relevant reference information with the relevant reference information as a starting point, and a means for processing the acquired relevant document information according to the processing conditions related to the reference information which becomes the starting point.



Data supplied from the esp@cenet database - Worldwide



【特許請求の範囲】

【請求項1】 参照情報と、再帰的取得の条件と、取得した文書の処理条件とを関連づけて記述した管理テーブルを保持する手段と、  
参照情報の指定を受ける手段と、  
当該指定された参照情報を起点として、当該参照情報で参照される文書データ内に設定されている参照情報に基づいて、当該起点となった参照情報に関連づけられた再帰的取得の条件に従って文書データを再帰的に取得する手段と、  
起点となった参照情報に関連づけられた処理条件に従って、当該取得した文書データを処理する手段と、  
を有することを特徴とする文書処理装置。

【請求項2】 参照情報と、再帰的取得の条件と、取得した文書の処理条件とを関連づけて記述した管理テーブルを保持する手段と、  
参照情報の指定を受ける手段と、  
当該指定された参照情報を起点として、当該参照情報で参照される文書データ内に設定されている参照情報に基づいて、当該起点となった参照情報に関連づけられた再帰的取得の条件に従って文書データを再帰的に取得する手段と、  
取得した文書データを特定する情報を提供し、取得した各文書データを処理対象として指定するか否かの設定を受け付ける手段と、  
起点となった参照情報に関連づけられた処理条件に従って、処理対象として指定された文書データを処理する手段と、  
を有することを特徴とする文書処理装置。

【請求項3】 請求項1又は2に記載の文書処理装置において、  
前記管理テーブルに記述された再帰的取得の条件は、条件式と、パラメータとからなることを特徴とする文書処理装置。

【請求項4】 請求項1から3のいずれかに記載の文書処理装置において、  
さらに、過去に取得した文書データをキャッシュして蓄積する手段を備え、  
前記再帰的取得の条件は、当該蓄積した文書データとの比較に基づく第1条件であることを特徴とする文書処理装置。

【請求項5】 請求項1から4のいずれかに記載の文書処理装置において、  
前記再帰的取得の条件は、  
参照情報で参照される他の文書データが現存しているか否かに基づく第2条件と、  
参照情報で参照される他の文書データが事前に指定されたキーワードを含むか否かに基づく第3条件と、  
参照情報で参照される他の文書データがプログラム起動に関する記述を含むか否かに基づく第4条件と、

参照情報により示される他の文書データの種別情報が事前に設定された種別情報に相当するか否かに基づく第5条件と、

参照情報で参照される他の文書データと、起点となった参照情報で参照される文書データとの電気通信網上の位置関係に基づく第6条件と、

のいずれかを含むことを特徴とする文書処理装置。

【請求項6】 請求項5に記載の文書処理装置において、  
前記再帰的取得の条件は、第1から第6条件のうち、複数の条件の組み合わせにより記述されることを特徴とする文書処理装置。

【請求項7】 請求項1から6のいずれかに記載の文書処理装置において、  
前記文書データを取得する手段が、所定の時間内に文書データが取得できないときには、当該文書データの取得の処理を中断することを特徴とする文書処理装置。

【請求項8】 請求項1から7のいずれかに記載の文書処理装置において、  
前記文書データを取得する手段が、取得の対象となった文書データのサイズを認識し、当該サイズが事前に設定されたサイズ上限を超える場合には、当該文書データの取得の処理を中断することを特徴とする文書処理装置。

【請求項9】 請求項1から8のいずれかに記載の文書処理装置において、  
前記文書データを取得する手段が、さらに、参照情報により参照される文書データがアクセス制限されている場合には、当該文書データの取得処理を中断することを特徴とする文書処理装置。

【請求項10】 請求項1から9のいずれかに記載の文書処理装置において、  
前記文書データを再帰的に取得する手段が、再帰的に取得した文書データを特定する情報を管理し、当該情報に基づいて参照情報のループを検出することを特徴とする文書処理装置。

【請求項11】 請求項10に記載の文書処理装置において、  
前記文書データを再帰的に取得する手段が、参照情報のループを検出したときには、ユーザに対し、ループの存在を報告することを特徴とする文書処理装置。

【請求項12】 請求項1から11のいずれかに記載の文書処理装置において、  
前記文書データを再帰的に取得する手段が、過去に取得した文書データを特定する情報を履歴情報として蓄積し、参照情報に従って文書データを取得する際に、当該履歴情報を参照し、当該履歴情報に基づいて文書データの取得を行うことを特徴とする文書処理装置。

【請求項13】 参照情報と、再帰的取得の条件と、取得した文書の処理条件とを関連づけて記述した管理テーブルを保持する手段と、

参照情報の指定を受ける手段と、当該指定された参照情報を起点として、当該起点となった参照情報で参照される文書データを取得して提供するとともに、当該提供した文書データに設定された参照情報を抽出して提供する手段と、前記抽出した参照情報の各々について、当該起点となった参照情報に関連づけられた再帰的取得の条件に従って文書データを再帰的に取得する手段と、起点となった参照情報に関連づけられた処理条件に従って、当該取得した文書データを処理する手段と、を有することを特徴とする文書処理装置。

【請求項14】 参照情報と、再帰的取得の条件と、取得した文書の処理条件とを関連づけて記述した管理テーブルを参照し、指定された参照情報を起点として、当該指定された参照情報で参照される文書データ内に設定されている参照情報に基づいて、当該起点となった参照情報に関連づけられた再帰的取得の条件に従って文書データを再帰的に取得する工程と、起点となった参照情報に関連づけられた処理条件に従って、当該取得した文書データを処理する工程と、を有することを特徴とする文書処理方法。

【請求項15】 参照情報と、再帰的取得の条件と、取得した文書の処理条件とを関連づけて記述した管理テーブルを参照し、指定された参照情報を起点として、当該参照情報で参照される文書データ内に設定されている参照情報に基づいて、当該起点となった参照情報に関連づけられた再帰的取得の条件に従って文書データを再帰的に取得する工程と、取得した文書データを特定する情報を提供し、取得した各文書データを処理対象として指定するか否かの設定を受け付ける工程と、起点となった参照情報に関連づけられた処理条件に従って、処理対象として指定された文書データを処理する工程と、を有することを特徴とする文書処理方法。

【請求項16】 参照情報と、再帰的取得の条件と、取得した文書の処理条件とを関連づけて記述した管理テーブルを保持するモジュールと、指定された参照情報を起点として、当該参照情報で参照される文書データ内に設定されている参照情報に基づいて、当該起点となった参照情報に関連づけられた再帰的取得の条件に従って文書データを再帰的に取得するモジュールと、起点となった参照情報に関連づけられた処理条件に従って、当該取得した文書データを処理するモジュールと、を有する文書処理プログラムを格納したことを特徴とするコンピュータ読み取り可能な記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、参照情報を含む文

書を処理する文書処理装置に関する。

【0002】

【従来の技術】インターネットの発展に伴って、パーソナルコンピュータ（PC）上のWebブラウザから参照情報としてのURL（Uniform Resource Locators）を指定して、Webページを表示させ、または印刷させることが広く行われている。ここで、Webページには、さらに他のWebページへのリンクが参照情報として設定されているのが一般的である。

【0003】インターネット上のWebページにおいては、この参照情報を用いて関連のあるWebページが相互にリンクされているのが普通であり、特定のWebページを参照しているうちに他の関連あるWebページも一緒に見たいとか、印刷時に一緒に手に入れたいという場合もある。例えば、天気予報のWebページを参照しているときに、東京地方の明日の天気掲載ページとともに、週間予報のページも一緒に印刷したい場合がある。

【0004】かかる場合のために、指定したURLで参照されるWebページに含まれる全てのリンク先のWebページを取得し、印刷を行う機能を備えたWebブラウザが存在する。この機能を利用することにより、関連のあるWebページを再帰的に取得して印刷処理することができる。

【0005】

【発明が解決しようとする課題】しかしながら、上記従来のWebブラウザでは、特定のWebページからリンクされた他のWebページの全てを取得して印刷することはできるものの、このようにするときわめて広範なページが取得され、印刷すると用紙が無駄になる場合がある。特に、いわゆるフルプリント（指定したURLで参照されるWebページを取得して印刷する印刷形態）を実装したプリンタ装置においては、取得したドキュメントをプレビューしてから印刷することができない（直接印刷されてしまう）ため、用紙の無駄を防止することが困難であった。

【0006】そこで、例えば特開平10-254912号公報「ワールドワイド・ウェブ・ニュース検索システム」では、事前に設定された複数のサイトからループをなすリンクを辿ることがないように文書データを取得してレイアウト処理を行う技術等が開示されているが、このワールドワイド・ウェブ・ニュース検索システムでは、リンクを辿る際の条件設定等が常にレイアウト情報の設定と混在して行われるので、設定内容が複雑になって利便性が低いうえ、フルプリントに配慮していないという問題点があった。

【0007】本発明は上記実情に鑑みて為されたもので、文書データの再帰的取得の条件設定を理解しやすくして利便性を向上でき、フルプリントに配慮した文書処理装置を提供することを目的とする。

【0008】

【課題を解決するための手段】上記従来例の問題点を解決するための本発明は、文書処理装置であって、参照情報と、再帰的取得の条件と、取得した文書の処理条件とを関連づけて記述した管理テーブルを保持する手段と、参照情報の指定を受ける手段と、当該指定された参照情報を起点として、当該参照情報で参照される文書データ内に設定されている参照情報に基づいて、当該起点となった参照情報に関連づけられた再帰的取得の条件に従って文書データを再帰的に取得する手段と、起点となった参照情報に関連づけられた処理条件に従って、当該取得した文書データを処理する手段と、を有することを特徴としている。これにより、レイアウトのような印刷の体裁に関する設定と関係なく再帰的取得の条件を可能とすることで利便性を向上できる。

【0009】また、上記従来例の問題点を解決するための本発明は、文書処理装置であって、参照情報と、再帰的取得の条件と、取得した文書の処理条件とを関連づけて記述した管理テーブルを保持する手段と、参照情報の指定を受ける手段と、当該指定された参照情報を起点として、当該参照情報で参照される文書データ内に設定されている参照情報に基づいて、当該起点となった参照情報に関連づけられた再帰的取得の条件に従って文書データを再帰的に取得する手段と、取得した文書データを特定する情報を提供し、取得した各文書データを処理対象として指定するか否かの設定を受け付ける手段と、起点となった参照情報に関連づけられた処理条件に従って、処理対象として指定された文書データを処理する手段と、を有することを特徴としている。

【0010】さらにここで、前記管理テーブルに記述された再帰的取得の条件は、条件式と、パラメータとからなることが好ましい。このようにすることで、各条件の設定が簡便になり、利便性を向上できる。

【0011】さらに、過去に取得した文書データをキャッシュして蓄積する手段を備え、前記再帰的取得の条件は、当該蓄積した文書データとの比較に基づく第1条件であることも好ましい。

【0012】さらに、前記再帰的取得の条件は、参照情報で参照される他の文書データが現存しているか否かに基づく第2条件と、参照情報で参照される他の文書データが事前に指定されたキーワードを含むか否かに基づく第3条件と、参照情報で参照される他の文書データがプログラム起動に関する記述を含むか否かに基づく第4条件と、参照情報により示される他の文書データの種別情報が事前に設定された種別情報に相当するか否かに基づく第5条件と、参照情報で参照される他の文書データと、起点となった参照情報で参照される文書データとの電気通信網上の位置関係に基づく第6条件と、のいずれかを含むことも好ましい。これらの条件により、プルプリントに配慮した条件設定が可能になる。さらに、前記

再帰的取得の条件は、第1から第5条件のうち、複数の条件の組み合わせにより記述されることも好ましい。

【0013】さらに、前記文書データを取得する手段が、所定の時間内に文書データが取得できないときには、当該文書データの取得の処理を中断することも好ましい。これにより、プルプリント等の際に、他の印刷処理に関係なく印刷処理を続行でき、プルプリントに適した条件設定を可能としている。また、当該文書データのサイズが事前に設定されたサイズ上限を超える場合に当該文書データの取得の処理を中断してもよい。

【0014】また、前記文書データを取得する手段が、さらに、参照情報により参照される文書データがアクセス制限されている場合には、当該文書データの取得処理を中断することも好適である。

【0015】また、前記文書データを再帰的に取得する手段が、再帰的に取得した文書データを特定する情報を管理し、当該情報に基づいて参照情報のループを検出することも好適である。

【0016】さらに、前記文書データを再帰的に取得する手段が、参照情報のループを検出したときには、ユーザに対し、ループの存在を報告する手段をさらに含むことも好適である。

【0017】また、前記文書データを再帰的に取得する手段が、過去に取得した文書データを特定する情報を履歴情報として蓄積し、参照情報に従って文書データを取得する際に、当該履歴情報を参照し、当該履歴情報に基づいて文書データの取得を行うことも好適である。

【0018】さらに、上記従来例の問題点を解決するための本発明は、文書処理装置において、参照情報と、再帰的取得の条件と、取得した文書の処理条件とを関連づけて記述した管理テーブルを保持する手段と、参照情報の指定を受ける手段と、当該指定された参照情報を起点として、当該起点となった参照情報で参照される文書データを取得して提供するとともに、当該提供した文書データに設定された参照情報を抽出して提供する手段と、前記抽出した参照情報の各々について、当該起点となった参照情報に関連づけられた再帰的取得の条件に従って文書データを再帰的に取得する手段と、起点となった参照情報に関連づけられた処理条件に従って、当該取得した文書データを処理する手段と、を有することを特徴としている。

【0019】さらに、上記従来例の問題点を解決するための本発明は、文書処理方法において、参照情報と、再帰的取得の条件と、取得した文書の処理条件とを関連づけて記述した管理テーブルを参照し、指定された参照情報を起点として、当該参照情報で参照される文書データ内に設定されている参照情報に基づいて、当該起点となった参照情報に関連づけられた再帰的取得の条件に従って文書データを再帰的に取得する工程と、起点となった参照情報に関連づけられた処理条件に従って、当該取得

した文書データを処理する工程と、を有することを特徴としている。

【0020】また、上記従来例の問題点を解決するための本発明は、文書処理方法であって、参照情報と、再帰的取得の条件と、取得した文書の処理条件とを関連づけて記述した管理テーブルを参照し、指定された参照情報を起点として、当該参照情報で参照される文書データ内に設定されている参照情報に基づいて、当該起点となった参照情報に関連づけられた再帰的取得の条件に従って文書データを再帰的に取得する工程と、取得した文書データを特定する情報を提供し、取得した各文書データを処理対象として指定するか否かの設定を受け付ける工程と、起点となった参照情報に関連づけられた処理条件に従って、処理対象として指定された文書データを処理する工程と、を有することを特徴としている。

【0021】

【発明の実施の形態】本発明の実施の形態について図面を参照しながら説明する。本発明の実施の形態に係る文書処理装置としてのプリンタ1は、図1に示すように、制御部11と、プリントエンジン12とから基本的に構成され、制御部11は、CPU21と、ROM22と、RAM23と、ハードディスク24と、制御インタフェース25と、ネットワークインタフェース26と、システムバス27と、ビデオインタフェース28と、操作表示パネル29と、システム時計部30とから構成されている。

【0022】CPU21は、ROM22やハードディスク24に格納されているプログラムに従い、システムバス27を介して制御部11の各部を制御する。ROM22は、CPU21によって処理されるプログラムを格納している。具体的にこのROM22には、ネットワークインタフェース26を介してデータを取得する処理のプログラム等が格納されている。RAM23は、CPU21のワークメモリとして動作し、一般的にリフレッシュ動作が不要で制御が簡便なSRAM (Static RAM) である。このRAM23には、CPU21からの指示によりプログラムの処理上必要となる変数やデータを格納する。

【0023】ハードディスク24は、CPU21が参照する再帰的取得の条件と、処理の条件とが図2に示すような態様で格納されている。本実施の形態において特徴的なことは、設定番号（設定No）の順に、処理条件としての直接印刷を行うか否かのフラグ（A）と、起点となる参照情報としてのURL（B）と、適用／除外フラグ（C）と、再帰的取得の条件ID（D）と、再帰的取得の条件に対するパラメータとしての条件値（E）とが関連づけられていることである。ここで、適用／除外フラグ（C）は、指定した再帰的取得の条件に合致するものを取得するか、条件に合致しないものを取得するかの設定を行うもので、いわば、再帰的取得の条件に対する

否定（NOT）を指定するものである。また、再帰的取得の条件ID（D）は、事前に設定された条件式に対する条件識別番号（識別子）であって、例えば「URLの構造的条件」のときには、IDが「1」となる。また、この構造的条件の場合のパラメータ（条件値）としては、起点のレベルと、例外の有無と、拡張子（ファイル種別）を利用したフィルタリングとが設定可能である。このように、IDと、条件式と、条件値との一例を表にまとめると、図3に示すようなものとなる。そして、条件値（E）がこの図3に示した条件値に対応してCPU21により解釈される。この図2のテーブルが本発明の管理テーブルに相当し、ハードディスク24が、このテーブルを保持する手段に相当する。図2に示したテーブルを用いて文書データの再帰的取得及び処理を行うCPU21の処理については後に詳しく説明する。

【0024】また、このハードディスク24には、CPU21が過去に取得した文書データに関する情報が図4に示すような履歴情報として蓄積されている。この履歴情報は、履歴番号（履歴No）で管理され、取得した文書データの参照情報（アクセスURL）と、当該URLへの最終アクセス日時と、アクセス回数とを記録したものである。

【0025】さらに、このハードディスク24には、CPU21によりユーザがブックマークとして登録したURLが図5に示すように、登録番号（登録No）で管理され、登録名（操作表示パネル29で確認する際に表示される文字列）と、登録したURLと、登録日時と、利用回数とが記録される。

【0026】さらに、このハードディスク24は、取得した文書データをキャッシュして保持している。また、このキャッシュを管理するためのテーブルが図6に示すように、処理番号（処理No）と、設定番号（設定No）と、取得番号（取得No）と、取得した文書データのURLと、取得日時と、選択の有無とが関連づけて格納されている。この処理番号は、キャッシュした際の処理ごとのIDであり、設定番号は、条件のIDに対応する番号であり、取得番号は、処理ごとの連番である。また、ユーザに提示してから選択する際の処理に配慮して選択の有無を表すフラグが設定可能となっている。これらの処理番号や設定番号などは、CPU21の処理の便宜として用いられる。

【0027】制御インタフェース25は、プリントエンジン12と接続され、CPU21から入力される指示に応じて、プリントエンジン12に対して制御信号を出力する。また、この制御インタフェース25は、プリントエンジン12から入力される制御に関する信号をCPU21に出力する。ネットワークインタフェース26は、電気通信網としてのインターネットに接続され、インターネットを介して取得したデータをCPU21に出力し、また、CPU21から入力されるデータをインター

ネットを介して送信する。

【0028】システムバス27は、制御部11の各部の間でデータを相互に伝送する。ビデオインタフェース28は、CPU21から入力される指示に基づいて、画像データをプリントエンジン12に出力する。操作表示パネル29は、例えば液晶タッチパネルや、LED、各種ボタン、スイッチなどからなり、ユーザに対して情報を提示し、また、ユーザが行った操作の内容をCPU21に出力する。システム時計部30は、時計用チップからなり、時刻の計時を行って、当該時計用チップが計時する時刻の情報（年月日、時分秒からなる情報）をCPU21に提供する。このシステム時計部30は、文書処理装置としてのプリンタ1の電源が切られているときや、停電時等にも時刻情報を維持できるよう、バックアップ電源を備えていることが好ましい。

【0029】ここで、本実施の形態のCPU21の処理動作について説明する。本実施の形態においては、本発明の参照情報の指定を受ける手段、文書データを再帰的に取得する手段、文書データを処理する手段、設定を受け付ける手段、キャッシュして蓄積する手段、参照情報を抽出して提供する手段の各々は、このCPU21におけるソフトウェア処理として実現されている。すなわちCPU21は、主として条件の設定処理と、文書データの取得処理と、文書データの出力処理とを行うもので、以下、これらについて順番に説明する。

【0030】まずCPU21は、操作表示パネル29から入力されるユーザからの指示により、条件の設定処理を開始し、図7に示すような設定画面を操作表示パネル29に提供する。設定番号は連番で既に設定されている条件を参照して次の順番を付してもよいし、削除された設定があって欠番が発生している場合には、欠番から任意に割り当てておくのもよい。また、ユーザが任意に設定できるようにしておくのも好適である。

【0031】ユーザは、この条件設定画面を利用して、取得処理の起点となるURL（参照情報）と、設定条件とを入力し、「OK」ボタンをプレスして設定を完了する。設定とともに、取得・印刷の処理を開始したい場合には、「PRINT」ボタンをプレスして即時実行させてもよい。また、設定を取りやめる場合には「CANCEL」ボタンをプレスする。ここで、起点となるURLの情報が図2のテーブルにおける指定URL（B）の設定に影響する。

【0032】この図7に示した設定画面の例では、再帰的取得の条件を「設定条件」欄内に設けられたラジオボタン要素や、プルダウンメニュー要素、及び文字列入力要素を操作して設定する。具体的に、「条件内容」のプルダウンメニュー要素には、図3に示した条件内容から選択できるようになっている。すなわち、例えばURLの構造関係に基づく条件を設定したい場合には、このプルダウンメニュー要素を操作して「URLの構造的条件」を選択し、条件値を対応する文字列入力要素に入力する。

件」を選択し、条件値を対応する文字列入力要素に入力する。

【0033】この条件内容のプルダウンメニュー要素では図6に示したキャッシュした文書データに基づく条件（本発明の第1条件）と、文書データが現存しているか否かに基づく条件（本発明の第2条件）と、指定したキーワードを含むか否かに基づく条件（本発明の第3条件）と、文書データの種別に関する条件（本発明の第5条件）と、電気通信網上の位置関係の一例としてのURLの構造的条件（本発明の第6条件）が少なくとも選択可能になっている。この設定が、図2に示したテーブルの条件ID（D）に影響する。また、併せて入力された条件値が図2の条件値（E）に影響する。

【0034】さらに、本実施の形態において特徴的なことは、このように指定した条件に合致するものを取得対象とするのか（適用条件とするか）、条件に合致しないものを取得対象とするのか（除外条件とするか）をラジオボタン要素により選択可能になっていることである。この適用／除外設定が、図2に示したテーブルの適用／除外フラグ（C）に影響する。

【0035】さらに、本実施の形態において特徴的なことは、直接印刷を行うか、取得処理の結果表示を行うかの選択をラジオボタン要素を利用して行うことができることである。この直接印刷の有無の設定が直接印刷フラグ（A）に影響する。

【0036】CPU21は、この設定画面においてユーザから入力された条件を図2に示すテーブルの形態で、ハードディスク24に蓄積する。

【0037】次に、CPU21の文書データの取得処理について説明する。CPU21は、図7に示した設定画面において「PRINT」ボタンのプレス操作を受けて、又は、指示により図8に示す特定のURLに対する取得処理を開始する。後者の場合には、事前にURLが入力され、処理開始の指示を受けて、図8の処理を開始することになる。

【0038】まず、CPU21は、指定された条件を満たす文書データを取得するために、起点として設定されたURLを取得対象URLの回数としてRAM23に格納し（S1）、図9に示す文書取得のサブルーチンをコールする（S2）。

【0039】CPU21は、文書取得のサブルーチンとして、RAM23に格納された取得対象URLから文書データを取得する要求をネットワークインタフェース26を介して送信する（S11）。この要求はHTTP（Hyper Text Transfer Protocol）等の文書データの転送が可能なプロトコルを利用して行われる。そして、CPU21は、タイマ（図示せず）を起動し（S12）、ネットワークインタフェース26により何らかのデータが取得されたか否かを調べる（S13）。ここで、データが取得されていなければ（Noならば）、処理S12

で起動したタイマがタイムアップしたか否かを調べ（S14）、タイムアップしていなければ（Noならば）処理S13に戻って処理を続け、タイムアップしていれば（Yesならば）処理を中断して、図8のメインルーチンに戻る（A）。

【0040】一方、CPU21は、処理S13において、何らかのデータが取得されたならば（Yesならば）、当該取得されたデータがエラー（例えばアクセス制限がされていることによるエラーや、文書データが現存しないために取得不能であることを知らせるエラー、又はトランザクションがタイムアウトしたことを知らせるエラー等）であるか否かを調べ（S15）、エラーであれば（Yesならば）処理を中断して図8のメインルーチンに戻る（A）。また、処理S15において、取得されたデータが文書データであれば（Noならば）、ハードディスク24にキャッシュして蓄積し、アクセスの記録を図4のテーブルに記録し、キャッシュの履歴を記述して図6のテーブルに蓄積する（キャッシュ等の記録処理を行う；S16）。そしてCPU21は、ハードディスク24にキャッシュした文書データを解析して文書データ内に設定されたURLを抽出し（S17）、抽出したURLのそれぞれに対し、起点として設定されたURLに対応する再帰的取得の条件に合致するか否かを調べ（S18）、合致しているURLがあれば（URLあり）、当該URLについて順次、図9の取得処理を再帰的に実行する（S19）。

【0041】そして、処理S18において、条件に合致するURLがなくなると（URLなし）、処理していない文書データがあるか否かを調べ（S20）、処理していない文書データがあれば（Yesならば）、当該文書データからURLを抽出する処理S17から繰り返して実行する。また、処理していない文書データがなければ（Noならば）、サブルーチンの処理を完了して、図8の処理に戻る。

【0042】ここで、条件に合致するか否かを調べる処理S18においては、図2に示すテーブルで設定された条件の内容が適用される。具体的に、URLの構造的条件が指定されているときには、起点としてのURLと、抽出されたURLとが比較され、条件値として設定された値に応じて条件に合致しているか否かが判断される。すなわち、条件値が「2階層下まで」を表している場合には、起点のURLよりも抽出したURLの階層が2階層下以内にあるときのみ、URLありとして、取得処理が再帰的に実行される。この比較は、一致部分を排除した後パスのデリミタ（一般的には半角の「/」（スラッシュ））の数をカウントして行うことができる。また、条件式として拡張子が設定されている場合には、抽出されたURLの拡張子が条件値に設定された拡張子と比較される。

【0043】尚、ここでは、文書データが現存しない場

合には処理を中断するとしたが、処理S15においてエラーの内容を解析し、当該エラーの内容が起点として設定されたURLに対応する取得処理の条件に相当する場合にのみ処理を中断することとし、取得処理の条件に相当しない場合には、処理S18に移行して未処理のURLの取得処理を行ってもよい。

【0044】さらに、処理S15においては、取得したデータが文書データであっても、当該文書データ内に記述されたものがCGI（Common Gateway Interface）プログラムとしてのperlスクリプトであったり、Java（商標）アプレットである場合など場合には、ユーザからの事前指示により取得を中断して処理S18に移行し、未処理のURLの取得処理を行うことが好ましい。このユーザからの事前指示が本発明のプログラム起動に関する記述である第4条件に相当する。

【0045】さらに、処理S19において、取得処理の対象となるURLに対し、既にアクセスしていないかどうかをCPU21が図4の態様でハードディスク24に格納した履歴情報に基づいて判断し、既にアクセスしている場合には、当該URLに対する2度目のアクセスを行わないようにしてもよい。同様に、図4の履歴情報の最終アクセス日時を参照して、これが所定期間を経過していない場合には、取得処理を行わずに次のURLの処理に移行することが好ましい。また、履歴情報ではなく、キャッシュ情報に基づく取得日時に基づいて、所定期間を経過していない場合には、取得処理を行わずに次のURLの処理に移行することとしてもよい。また、アクセス履歴に基づきループが形成されていることが検出されると、次のURLの処理に移行してもよいし、この場合にユーザに対してループがあったことを報知するのでも好ましい。

【0046】一方、図8に示すメインルーチンでは、図9のサブルーチンから処理が戻ると（A）、起点としてのURLに対応する処理条件に即して、直接印刷のフラグが「1」であるか否かが判断され（S3）、直接印刷のフラグが「0」であれば（直接印刷しない場合：Noならば）、履歴としてハードディスク24に蓄積した図4のテーブルに基づき、図10に示すようなインタフェース画面が操作表示パネル29に提供される（S4）。

【0047】この図10のインタフェース画面では、取得履歴に基づくURLが対応するラジオボタン要素とともに一覧表示され、印刷したい文書をラジオボタン要素の操作により選択して「PRINT」ボタンをプレスして、印刷を行うことになる。

【0048】この一覧表示からURLが選択され「PRINT」ボタンがプレスされると、当該対応する文書データに基づく画像形成処理が行われ、ビデオインタフェース28を介してプリントエンジン12に出力され、制御インタフェース25を介してプリントエンジン12が制御されて、プリントエンジン12で、指示されたUR



しに対応する文書データが印刷出力される（S5）。この処理S5が出力の処理に対応する。

【0049】これにより、簡便な設定に基づき、必要なURLに存在する文書データのみが印刷出力されることになる。

【0050】さらにユーザは、「ブックマーク」の条件式を設定することで、図9の処理S18においてCPU21に対し、図5に示したブックマークされたURLであるか否かにより文書データの取得処理を行わせることもできる。この場合には、図5の利用回数を参照し、利用回数が所定の値を超えているもののみ、又は所定の値に達していないものに限定して取得処理を行わせることもできる。さらに登録日時に応じた条件としてもよい。また、「履歴」の条件式を設定すれば、CPU21は、処理S18において図4のアクセスURLを参照して取得処理を行うことになる。また、図7に示した設定画面において、適用／除外フラグの設定に応じて、CPU21は、例えばブックマークされていないURLのみを取得対象とする。また、CPU21は、処理S18において取得しようとしている文書データのファイルサイズが所定のサイズ（例えば事前に設定されたサイズ上限）よりも大きいときには当該文書データの取得処理をスキップして次のURLに対する処理に移行してもよい。この場合には、ユーザに対し、その旨を報知することも好ましい。

【0051】さらにここまでの説明においては、図7に示したように、条件式及びそれに付随する条件値は単独で設定することとしているが、これらの条件式や条件値で指定される再帰的取得条件同士のAND（論理積）やOR（論理和）、XOR（排他的論理和）等の論理演算の結果に基づく再帰的取得条件を設定可能としてもよい。

【0052】さらに、図9において、文書の取得処理を中断したときには、印刷の処理が行われないようになっているが、図9の（A）から図8の（B）に示した処理に移行するようにして、そこまでの処理で取得した文書データの印刷を行わせることとしてもよい。このようにすれば、エラーが検出されるまでの文書データが印刷できる。

【0053】また、ここまでの説明においては、文書処理装置がプリンタのように操作表示パネルが比較的小さい場合を例としていたが、本発明の文書処理装置としては、一般的なWebブラウザソフトウェアに対するブラウザインプログラムの様な形態で実装されていてもよい。この場合には、アクセス中の文書データに設定された他の文書データへの参照情報を条件に従って再帰的に取得することになる。ここで条件は、既に説明したように、ブラウザのブックマークによってもよいし、ブラウザが蓄積したキャッシュの情報や、過去のアクセス履歴

によってもよい。また、参照情報により参照される文書データが現存しない場合（ブローンリンクとなっている場合）には、取得を中断する等の条件としてもよい。

【0054】このように、本実施の形態によれば、文書データが現存しない場合などにも対応しつつ、所定の再帰的取得条件を満たす文書データのみが選択的に取得されて、当該条件に関連する処理条件に従い、文書処理が行われるので、設定が容易で利便性が高く、文書データを効率的に取得できるなど、フルプリントに配慮した取得及び処理を行うことができる。

【0055】

【発明の効果】本発明によれば、参照情報と、再帰的取得の条件と、取得した文書の処理条件とを関連づけて記述した管理テーブルを保持し、参照情報の指定を受けて、当該指定された参照情報を起点として、当該参照情報で参照される文書データ内に設定されている参照情報に基づいて、当該起点となった参照情報に関連づけられた再帰的取得の条件に従って文書データを再帰的に取得し、当該取得した文書データを処理する文書処理装置としているので、設定が容易で利便性が高く、文書データを効率的に取得できるなど、フルプリントに配慮した取得及び処理を行うことができる。

【図面の簡単な説明】

【図1】 本発明の実施の形態に係る文書処理装置としてのプリンタ1の構成ブロック図である。

【図2】 管理テーブルの一例を表す説明図である。

【図3】 設定可能条件の一例を表す説明図である。

【図4】 CPU21が管理する履歴情報のテーブルを表す説明図である。

【図5】 CPU21が管理するブックマークのテーブルを表す説明図である。

【図6】 CPU21が管理する取得履歴の一例を表す説明図である。

【図7】 再帰的取得条件の設定画面を表す説明図である。

【図8】 CPU21の処理の一例を表すフローチャート図である。

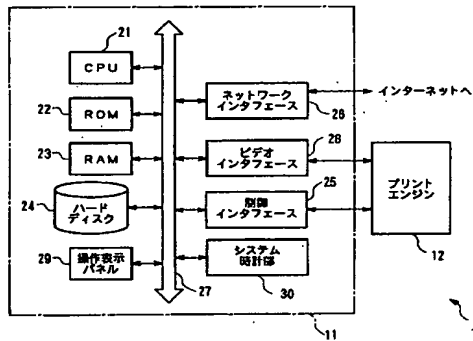
【図9】 CPU21の文書データの取得処理の一例を表すフローチャート図である。

【図10】 文書データの選択画面の一例を表す説明図である。

【符号の説明】

1 プリンタ、11 制御部、12 プリントエンジン、21 CPU、22 ROM、23 RAM、24 ハードディスク、25 制御インタフェース、26 ネットワークインタフェース、27 システムバス、28 ビデオインタフェース、29 操作表示パネル、30 システム時計部。

【図1】



【図2】

	A	B	C	D	E
条件 ID	条件 内容	指定URL	適用/ 除外	条件 値	条件値1 条件値2 条件値3
1	0	http://www.aaa.co.jp/index.html	適用	1	0 なし 未設定
2	1	http://www.aaa.co.jp/s/b/index.html	適用	1	2 なし 未設定
3	1	http://www.bbb.co.jp/abc/index.html	適用	2	一致 なし 未設定
4	0	http://www.bbb.co.jp/abc/index.html	適用	2	機下 あり jpeg
5	0	http://www.bbb.co.jp/abc/index.html	適用	3	一致 なし 未設定
6	0	http://www.bbb.co.jp/abc/index.html	適用	3	機下 なし 未設定
7	0	http://www.bbb.co.jp/s/index.html	適用	4	html, jpeg (未使用) (未使用)
8	0	http://www.aaa.co.jp/index.html	適用	1	0 あり jpeg
9	0	http://www.bbb.co.jp/s/index.html	除外	4	jpeg, leo (未使用) (未使用)

【図3】

ID	条件式	条件値1	条件値2	条件値3
1	URLの 機能的条件	起点レベル	例外の有無	拡張子
2	展 示	一致/機能的 条件	例外の有無	拡張子
3	ブックマーク	一致/機能的 条件	例外の有無	拡張子
4	拡張子一致	拡張子	(未使用)	(未使用)
		!		

【図7】

設定 No:

URL:

設定条件

☐ 適用条件 ☐ 除外条件

条件内容:

条件値1:

条件値2:

条件値3:

☐ 拡張子表示 ☐ 印刷

OK CANCEL PRINT

( 0 ) 101-350745 ( P2001-35T18

【図4】

履歴 No	アクセス URL	最終アクセス日時	アクセス回数
1	<a href="http://www.aaa.co.jp/index.html">http://www.aaa.co.jp/index.html</a>	1999/11/11 11:11	1
2	<a href="http://www.bbb.co.jp/abc/index.html">http://www.bbb.co.jp/abc/index.html</a>	1999/11/12 17:17	1
3	<a href="http://www.bbb.co.jp/abc/index.html">http://www.bbb.co.jp/abc/index.html</a>	1999/11/13 10:51	2

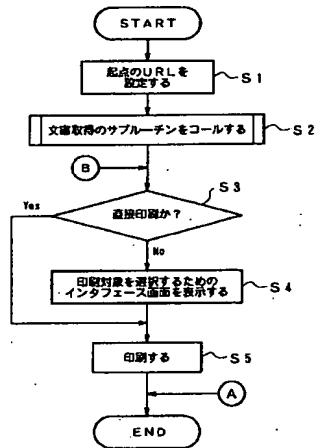
【図5】

登録 No	登録名	登録 URL	登録日時	利用回数
1	aaa	<a href="http://www.aaa.co.jp/index.html">http://www.aaa.co.jp/index.html</a>	1999/11/11 11:11	1
2	bbb1	<a href="http://www.bbb.co.jp/index.html">http://www.bbb.co.jp/index.html</a>	1999/11/12 17:17	1
3	bbb2	<a href="http://www.bbb.co.jp/abc/index.html">http://www.bbb.co.jp/abc/index.html</a>	1999/11/13 10:51	2

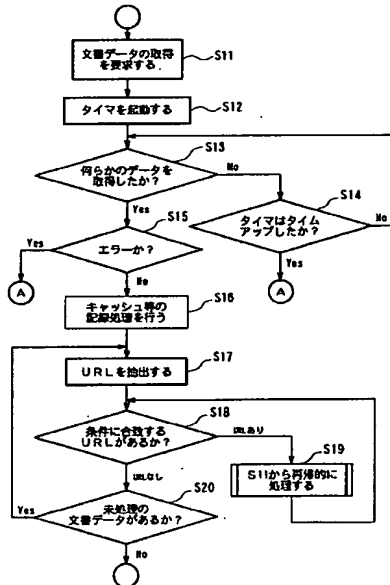
【図6】

処理 No	設定 No	取得 No	取得 URL	取得日時	選択
1	1	1	<a href="http://www.aaa.co.jp/index.html">http://www.aaa.co.jp/index.html</a>	1999/11/11 11:11	1
1	1	2	<a href="http://www.aaa.co.jp/aaa/aaa1.html">http://www.aaa.co.jp/aaa/aaa1.html</a>	1999/11/11 11:11	1
1	1	3	<a href="http://www.aaa.co.jp/aaa/aaa2.html">http://www.aaa.co.jp/aaa/aaa2.html</a>	1999/11/11 11:12	0
1	1	4	<a href="http://www.aaa.co.jp/aaa/aaa3.html">http://www.aaa.co.jp/aaa/aaa3.html</a>	1999/11/11 11:12	1
2	1	1	<a href="http://www.bbb.co.jp/index.html">http://www.bbb.co.jp/index.html</a>	1999/11/13 10:51	0
2	1	2	<a href="http://www.bbb.co.jp/bbb/bbb1.html">http://www.bbb.co.jp/bbb/bbb1.html</a>	1999/11/13 10:51	0
2	1	3	<a href="http://www.bbb.co.jp/bbb/bbb2.html">http://www.bbb.co.jp/bbb/bbb2.html</a>	1999/11/13 10:51	0

【図8】



【図9】



【図10】

設定 No: 1      処理 No: 1

取得文書:

- ☒ http://www.aaa.co.jp/index.html
- ☐ http://www.aaa.co.jp/aaa/aaa1.html
- ☐ http://www.aaa.co.jp/aaa/aaa2.html
- ☒ http://www.aaa.co.jp/aaa/aaa3.html

OK      CANCEL      PRINT